Innovation and social transmission in experimental micro-societies

McGuigan, Nicola; Burdett, Emily; Burgess, Vanessa; Dean, Lewis; Lucas, Amanda; Vale, Gillian; Whiten, Andrew

Published in:
Philosophical Transactions of the Royal Society B: Biological Sciences

DOI:
10.1098/rstb.2016.0425

Published: 05/12/2017

Document Version
Peer reviewed version

Innovation and social transmission in experimental micro-societies: exploring the scope of cumulative culture in young children

Nicola McGuigan[1,2], Emily Burdett[1,3,4], Vanessa Burgess[1], Lewis Dean[1], Amanda Lucas[1,5], Gillian Vale[1,6,7] and Andrew Whiten[1]

[1]Centre for Social Learning and Cognitive Evolution, School of Psychology & Neuroscience, University of St Andrews, South Street, St Andrews, KY16 9JP

[2]School of Life Sciences, Heriot-Watt University, EH14 4AS

[3]Institute of Cognitive and Evolutionary Anthropology, University of Oxford, OX2 6PN

[4]Brain, Belief, and Behavior Lab; Centre for Psychology, Behaviour, and Achievement, Coventry University, Coventry CV1 5FB, UK

[5]College of Life and Environmental Sciences, University of Exeter, TR10 9EZ

[6]National Center for Chimpanzee Care, Michale E. Keeling Center for Comparative Medicine and Research, The University of Texas MD Anderson Cancer Center, Bastrop, TX, USA

[7]Departments of Psychology and Philosophy, Neuroscience, Institute and Language Research Center, Georgia State University, Atlanta, GA, USA

Subject Areas:

cognition, evolution, behaviour

Keywords:

innovation, social learning, culture, cumulative culture, cultural evolution, tool-use

Author for correspondence:

Andrew Whiten

a.whiten@st-andrews.ac.uk

The experimental study of cumulative culture and the innovations essential to it is a young science, with child studies so rare that the scope of cumulative cultural capacities in childhood remains largely unknown. Here we report a new experimental approach to the inherent complexity of these phenomena. Groups of 3-4 year old children were presented with an elaborate array of challenges affording the potential cumulative development of a variety of techniques to gain increasingly attractive rewards. In contrast to a prior study, we found evidence for elementary forms of cumulative cultural progress, with inventions of solutions at lower levels spreading to become shared innovations, and some children then building on these to create more advanced but more rewarding innovations. This contrasted with markedly more constrained progress when children worked only by themselves, or if groups faced only the highest level challenges from the start. Further experiments that introduced higher-level inventions via the inclusion of older children, or that created ecological change, with the easiest habitual solutions no longer possible, encouraged higher levels of cumulative innovation. Our results show children are not merely ‘cultural sponges’, but when acting in groups, display the beginnings of cycles of innovation and observational learning that sustain cumulative progress in problem solving.

1. Introduction

The present century has seen a substantial blossoming of systematic research on cultural evolution in our own supremely cultural species [1-4]. Extensive evidence has similarly accumulated for social learning and cultural traditions in many vertebrate and even invertebrate species [1, 4, 5-8]: culture, in the sense of socially transmitted tradition, has been shown to be widespread in the animal kingdom [4].

The existence of such traditions depends upon two very different but complementary processes: first, some behavioural novelty must be created; then it must be repeatedly transmitted, diffusing amongst one or more populations [4, 9-11]. Legare and Nielsen [12] called these two contrasting processes – novel invention, and copying others – the ‘dual engines’ of culture.

These dual engines become yet more significant when we focus on the *cumulative culture* that many authors argue distinguishes our own from other cultural species [2, 3, 13-15]. Cumulative culture, the focus of the present paper, involves repeated cycles of invention and copying; later novelties are built, often repeatedly, on the shoulders of earlier ones, permitting ever more sophisticated cultural products to evolve [1-3, 12-18].

Given the significance of cumulative culture it is somewhat surprising that although varied descriptive analyses were published on the topic through the last century [19-21], systematic experimental studies emerged only in the last decade, remaining particularly rare in child research. The latter appears puzzling given that (i) childhood is generally regarded as a period of massive cultural assimilation [12, 22]; and (ii) one might expect great scope for the other side of the ‘dual engine’ – invention - as much of childhood is devoted to creative play [23, 24]. We briefly review the recent corpus of cumulative culture experiments in section 3 below as background to the child study we report in this paper. First, a core terminological issue should be clarified.

2. A note on terminology

Legare and Nielsen [12] refer to the first of their two engines of culture as ‘innovation’, as do many other authors [22]. However, it has become common to apply more demanding definitions that also bring in the other, ‘copying’ engine, for example defining the process of behavioural innovation as one that ‘results in new or modified learned behaviour and that introduces novel behavioural variants into a population’s repertoire' [9,10]. We adhere to this definition here, in line with the introduction to this journal issue [11]. This population-level definition of innovation makes it a core feature of culture, because to be counted as an innovation, a behaviour must 'diffuse through some segment of a group of potential users' [11]. It must therefore arise by a process we here call 'invention', possibly in the behaviour of a single individual, then adopted by one or more others to become an innovation. Elsewhere such an invention is itself often called an innovation [12, 22], so it is important to clarify our usage.

3. The experimental investigation of culture and the neglect of innovation

Cultural transmission has been subjected to varied forms of observational, statistical and experimental investigation, producing a wealth of discoveries [1-7, 12, 14, 15, 22, 25]. Controlled experiments offer substantial power in identifying social learning, but at the expense of the invention and innovation components. This is because in these 'diffusion' experiments, alternative behaviours are seeded in experimenter-trained initial models, and the differential spread of these across groups is tracked [24-28]. 'Asocial' control groups with no model may be included. In one approach, the 'transmission chain', a second individual watches the initial model, then becomes the model for a third, and so on along a chain, a method applied increasingly with both human and non-human participants, including children [28-30]. Alternatively, in 'open group diffusion' designs, models are introduced into whole groups, so that it is 'open' who watches and who learns from what they see [26,31]. This approach incorporates group dynamics that may be important in real-world cultural transmission.

However, the focus of such studies is typically social learning and what form the potential diffusion takes; the key initial 'innovation' element is provided by the experimenters, not participants. In the present study we sought to redress this balance by allowing the child participants to be free to generate inventions for task solution rather than initiating them by experimenter intervention. We then tracked evidence for any social transmission that turned these into shared innovations.

4. The experimental investigation of cumulative culture: innovation and transmission

Cumulative culture begs this more symmetric approach to innovation and social learning, because sequential steps in innovation have to come from the participants. Yet, despite its significance for human culture, experimental investigation of the cumulative aspect began only recently in humans [32] and in chimpanzees [33]. The human studies began with ‘laboratory micro-cultures’ such as building spaghetti-and-plasticine towers as high as possible [34, 35], later investigating the social learning processes necessary for transmission [35-37] and factors moderating cultural fidelity, conformity and diversity [38-42]. Allied approaches have been applied to the evolution of artificial ‘mini-languages’ in humans and non-humans [43, 44], although here the ‘cumulative’ effect is typically not the creation of more complex outcomes, but a regularization of the transmitted ‘syntax’ that is increasingly memorable and replicable. Finding children in diffusion chains progressively omitting inessential aspects of behaviour witnessed [45] appears ‘cumulative’ in only this perhaps marginal sense.

We are aware of only four pioneering experiments directly addressing cumulative culture with young children. First, having found that young chimpanzees failed to show cumulative social learning of an elaborate tool-use step that built on an earlier, simpler technique to exploit an artificial foraging device [33], Whiten et al. [46] offered the same options to 3-4 year old children. A majority of the children did show social learning of the second and more productive step, contrasting with the conservatism of the chimpanzees in the earlier study. However this study did not investigate further diffusion of the cumulative step achieved.

By contrast, Dean et al. [47] offered a three-stage puzzle box to young children, chimpanzees and capuchin monkeys in a group context, thereby creating an open group diffusion experiment. Increasingly attractive rewards could be obtained by a series of three steps of increasing difficulty, to instantiate accumulation. First, a door was simply slid to one side; second, it could be moved further only after pressing a tab to release it; and finally the door could be moved to reveal the best reward by turning a dial. Results showed that the children more often completed higher levels of the task than the non-human primates. The authors ascribed this to a suite of behavioural differences, including children’s greater tendency to replicate the approach of successful others, to share rewards gained, and to teach. However one of the most significant aspects of true cumulative cultural evolution is that the build-up of prior steps *enables* the further achievement of the higher ones. That could not be addressed given the apparatus design in this study, in which each higher step had to be enabled by a lower one. We do not know if similar species differences would have emerged if (i) participants were faced with step 3 only, or (ii) if all three steps had been simultaneous options. No asocial control condition was included to thus test if cumulative culture had indeed occurred at the group level. Our present study did include such a control, to address this core aspect of cumulative culture more directly.

A third study [48] used a diffusion chain design with five 'generations' in which young children were set a task of shifting as much rice from one container to another, with tool options ranging from a strip of card (least efficient) to a scoop and a bowl (most efficient). Some chains were seeded by an adult using the inefficient card, and others were unseeded. All seeded chains began by copying the card-use and half continued with this throughout, others switching to scoop-use. Results revealed little evidence for cumulative culture; rather, a few children gave up the card in favour of an option visibly better, the scoop. Seeded children's adoption and frequent persistence with the card option offers a potent illustration of the fidelity of child copying proclivities, and perhaps even 'over-imitation', a tendency to copy sub-optimal techniques when intentionally demonstrated by an adult [49]. The non-seeded chains showed a similar consistency and did not accumulate innovations. Just two children used the scoop and more efficient bowl in the final generation, but these putative instances of cumulation were achieved by just two of the 80 children in the experiment.

In the fourth and most recent study, researchers modified the 'spaghetti tower' paradigm earlier developed for diffusion studies with adults [34]. They first showed that creating a three-stick-high tower was beyond the young child sample and so described it as a 'culture-dependent product'. They then showed that young children would typically be able to copy such a construction by an adult (with or without sight of the actions needed to do so) [50]. This meets one criterion many writers on cumulative culture insist on, that the cumulative step socially learned must be one that is otherwise beyond learners [13, 15]. However the study did not demonstrate cumulative culture amongst children insofar as the children did not create the innovation copied, nor was its potential to diffuse amongst children demonstrated. Indeed, the framing of the study as novel in showing copying of a 'cumulative technological design' or 'culture-dependent product' is premature insofar as the experimental design is simply that typical of basic social learning studies in which participants in an asocial control condition fail, and those in a social learning condition succeed, whether in children [51] or chimpanzees [52]. We already know that children socially learn all manner of things from their surrounding culture they would otherwise not invent, but that leaves unanswered the question of the emergence of the broader package of cumulative culture in childhood.

There is thus little compelling evidence from these initial studies that the twin engines of cumulative cultural evolution are beginning to operate in a synergistic fashion in young children. Accordingly in the present study we initiated a broader based exploration of the scope of cumulative culture in child communities, building on the foundations provided by these earlier studies, but addressing their limitations. We note that the physical experimental arenas employed in the existing studies were very 'stripped-down'. In the real world, cumulative cultural change occurs in complex environments that afford multiple options wherever potential change or progress may occur. Cultural evolution is often tree-like, resembling phylogenetic branching in biological evolution, expressing both cumulative change and diversification, as in regionally differing cultural developments [1, 20, 54]. Accordingly we exposed children in different nursery classes to a complex 'model world' affording multiple behavioural options that could potentially generate these critical effects.

5. Overview and rationale of the study

In studies with chimpanzees (to be reported elsewhere), our approach was modelled upon Köhler's famous experiments [55] in which varied tool options were available to secure out-of-reach food rewards, so we called this a 'Köhler World'. For children, we created a more practicable 'table-top' version we called the 'Small World' (SW: figure 1), applying this in studies of both children and chimpanzees (the latter to be reported elsewhere). In its basic form the SW includes four levels of manipulanda that offer increasingly valued rewards requiring increasingly difficult means of attainment, such that higher levels may be enabled by cumulative progress through the lower ones. The first level can be achieved by manual (finger-poking) actions; the second is more demanding in that it requires use of a simple stick-tool; the third needs the fashioning of a more complex, longer tool and the fourth more complex tools yet, including adding a hook element, known to be challenging for young children [51]. At each level, there are four different types of exit manipulation that will release rewards, with each type recurring across the four different levels such that success on each higher level can potentially build on prior learning at a lower level. Additionally, multiple tool components can be combined and used in different ways at each juncture. The array is thus intentionally complex in these respects compared to prior experimental studies of cumulative culture. With this SW we are able to address questions about both cumulative progress (proceeding from lower to upper levels, requiring increasingly complex tools) and diversification (specialization by different children and groups).

We approached this investigation in the spirit of an ethological field study, designed to discover whether and to what extent the synergy between invention, innovation and copying that constitutes cumulative culture is already emerging in early childhood. Our ignorance on this question makes it inappropriate to hypothesise either that cumulative culture *would* be manifested, or that it would *not*, but in practice our experimental design and statistical analyses were directed at testing a series of predictions arising from the potential emerging operation, or not, of such cumulative culture at the ages examined.

Of course, that childhood is a period of great cultural learning is well established [12, 30, 45, 46, 49, 53]. Yet, the ontogeny of the larger capacity for cumulative culture, with its dependence on integration of invention and innovation, remains unknown. Reasons to predict evidence of cumulative culture in early development include this capacity being critical to our species' success [3, 13, 14], and also that play, often highlighted as a major source of creativity and innovation [23, 24] suffuses childhood at the stage we investigated. Reasons to predict the opposite, a lack of evidence for cumulative culture, include young children's remarkably unimaginative problem solving requiring innovative tool use [51, 54, 57], as when even six-year-olds fail to appreciate that a pipe-cleaner can be bent to make a useful hook tool (whereas they will promptly do this once they see it demonstrated) [51]. Similar limitations have been documented in using water to gain a floating prize [58], or a loop of cord to rake an object closer [13]. We know surprisingly little about the ontogeny of humanity's prodigious capacity for cumulative culture, and our study sought to make inroads into this largely uncharted territory.

We have structured the statistical analysis of our results in terms of four broad predictions consistent with the existence of childhood cumulative culture, each of these four being further delineated through a series of more specific predictions, as set out in supplementary tables S1 and S2. The first prediction was that the progress of groups of children, and also of individual children, would tend to display cumulative progress up the levels of the SW. The SW was specifically designed to offer the possibility of such cumulative progress from more elementary levels through increasingly challenging ones, with the higher steps potentially building on the lower ones. For example, what was learned about positioning rewards for removal using one's fingers at L1 could in principle be applied to moving them using the more challenging tools at higher levels; and what was learned about how a particular type of exit worked could be exploited when it was met again at higher levels, but requiring skill build-up because more complex forms of tool use were demanded. This first prediction of cumulative 'upwards' progress does not in itself yet identify this as a cultural process dependent on social learning, but represents one component of this larger phenomenon.

The second prediction was that groups would make greater progress up levels compared to children in two control conditions: an asocial control condition where children worked by themselves, and a further group condition in which only the most challenging level was available. Again this does not necessarily identify social learning as such, although social learning would of course facilitate such a contrast. Instead it could be that other more general social factors, such as confidence in working in groups, could create such an effect, but this would nevertheless be an interesting outcome, and could perhaps even be regarded as marginal or precursor manifestation of a cumulative culture.

The third broad prediction was that social learning was indeed at work, and possibly ubiquitous. This was tested by examining the extent to which a series of actions that children first successfully took, such as their first success at a particular exit, was biased by what they had witnessed beforehand.

The fourth and final prediction was focussed on the criterion for cumulative culture that progress up higher levels could be made by children different to those who had provided the innovation built on at lower levels, with such innovations then spreading further to others.

These four predictions were first addressed in a 'Phase A' of the study in which nine nursery classes of 3-4 year old children were presented with the SW over a series of sessions described below. In a ‘Phase B’ we then conducted two additional manipulations, each within groups who participated in Phase A, and predicted to enhance any evidence for cumulative culture. In one, exploring the extent to which cumulation may depend on occasional particularly gifted innovators, we created the appearance on the scene of a such an individual in the guise of an older child (Phase B (a)). In a second manipulation we created a significant ecological change of the kind thought to have possibly driven human cultural evolution in the past, in this case removing some of the lower levels of the SW, creating greater pressure for higher cumulative achievements (Phase B (b)).

6. Discovery, innovation and cumulative learning in child micro-societies: Materials and methods(Phase A)

(a) Participants

A total of 152 children (72 girls) with a mean age of 47 months (s.d. = 6 months) participated in experiments replicated in each of nine local nursery groups in Scotland (*M* number of children per group = 17, range = 8-25). An additional group of 15 participants (7 girls; *M* age = 45 months, s.d. = 5 months) took part in a control condition where only the top SW level was available, with a further 23 children (11 girls; *M* age = 46 months, s.d. = 5 months) participating individually in an asocial control condition. Demographic details are in supplementary table S3.

(b) Apparatus: the ‘Small World’

The ‘Small World’ (SW) was accommodated in a large box (figure 1; supplementary figure S1) from which four types of reward, ranging from a sticker to a small toy figure, contained in four differently coloured plastic eggs, and established to have increasing value to children (see supplementary methods) could be obtained by actions with four correspondingly escalating levels of difficulty. At Level 1 (L1), the reward (a paper sticker inside a plastic egg) could be manoeuvred from its central position in a box-like shelf (figure 1a) to one of four potential exit points by poking a finger in a horizontal slit and manoeuvring the egg along. Direct access to box shelves was prevented by plastic sheets and mesh screens on the front of the SW (supplementary figure S1 shows a photo). From each exit point a released egg would drop to a sloped ramp and roll out at the front of the SW (at one exit it instead came out at the front). Each of the four exits required a different procedure for egg release; the Ramp (R), Trapdoor (T), Pushdoor (P) and Lift (L) (figure 1b,c). At Level 2 (L2) the box-shelf was situated 10 cm from the front slit, so use of one of a variety of stick tools available in a pocket on the front of the SW was necessary to move the egg and operate an exit. Exits took essentially the same forms as at L1 (R,T,L,P) but now required tool use to open them. Such tools *could* be used at L1 also. Level 3 (L3) was a further 20 cm back so here the creation of longer tools became necessary for task solution. Options were available in two modes (figure 1d): *Combine*, in which some sticks could be inserted into the ends of thicker ones; and *Unfold*, where hinged links allowed initially folded tools to be unfolded. Contrasting colours helped to mark the Combine (yellow) and Unfold (green) families of tools (figure 1d). Level 4 (L4) was situated at the same distance as L3, but exits could be operated only by creating a 'hook' element in the different types of tool (through combining or unfolding). For further details see figure 1 and supplementary methods.

Unlike the apparatus of Dean et al. [44], the SW also made it possible *in principle* to attain higher levels without approaching them cumulatively. Importantly, this left open empirical questions as to whether or not evidence of cumulative progress in social groups would occur, and whether children tested individually in an 'asocial' condition, and others tested in groups but faced only with the highest level to attempt, would show that it was cumulation in groups that permitted higher levels to be attained.

(c) Procedure

In the principal 'open diffusion' *Group* condition, the SW was presented to nine nursery groups of children in their normal play areas, for one hour on four consecutive days. We opted for this open diffusion design, rather than an alternative such as a transmission chain of children (A-B, B-C, C-D etc.) because the children's capacity for cumulative progress in a context like the SW was so unknown. In a transmission chain, any child insufficiently motivated or competent might destroy any prospects for cumulation, whereas in an open diffusion context, there was the possibility of synergy between the most motivated and competent children to generate cumulative effects. We address the prospects for what alternative designs may in future add to such research, in the Discussion section below.

Introductory remarks to the children were intentionally minimal, restricted to the experimenters, already familiar to the children, saying “We’ve brought a big puzzle box from which you can try to get prizes”. Experimenters showed children the different reward items inside each of the four differently coloured eggs, adding that “Everyone can have a go, and you can do anything you like to get the prizes”. Whenever experimenters or nursery supervisors were asked questions, intentionally neutral pre-prepared responses were offered, such as “It’s rather tricky but you are doing really well” (see supplementary methods for details). No attention was drawn to the tools available in trays on the front of the SW.

At the start of each daily session an egg containing a reward was available in the centre of each box-shelf. For the first half-session any egg removed by children was immediately replaced by an experimenter from behind the SW to maintain overall motivation to participate. Subsequently, to avoid children focusing only on lower and easier levels, the re-bait period was progressively increased by 30 seconds and later by one minute in sessions one - three, until by session three it was four minutes, and in session four it was five minutes. Once lower-level rewards had been extracted, there was thus a reason for children to attempt higher levels, in addition to the greater value of the reward they could provide.

In an *asocial* condition, individual children were presented with the SW in a room off the main nursery area, with instructions matching those of the group condition. To encourage these children to attempt all levels if they wished to, rebating occurred at intervals of five minutes. The first session was limited to 15 minutes (judged an optimal period for the solo child involved), with children who successfully obtained a reward going on to two further 15-minute sessions on subsequent days.

A further control condition presented only L4 to one nursery group (L4C) for four hourly sessions, as a check to discover if children could succeed at this without the potential cumulative process that might occur in the nine main groups presented with L1-L4.

A video camera on the experimenter’s side of the SW recorded child actions on the SW, plus observations of these by nearby children. Video records were subsequently coded, with codings later checked by a second coder. ‘Observation’ was coded whenever a child was within 1 metre of the apparatus and oriented in the direction of the action. A second supplementary camera with a side view was used to clarify any codings where necessary. The full coding scheme is described in supplementary methods.

In the analyses described below, an ‘invention’ was defined minimally as the first success by a child at a specific exit at a specific level. Results were analysed to answer a series of questions about such inventions, their sharing by children to become ‘innovations’, the scope of any cumulative progress across levels in the SW, and the extent to which this relied on the social learning that would make it a case of simple cumulative culture. An overview of the main predictions and corresponding sub-predictions for Phase A is provided in supplementary table 1.

7. Discovery, innovation and cumulative learning in micro-societies: Phase A Results

(a) Testing the prediction that cumulative progress tends upwards across levels

(i) Was progress cumulative across sessions? Dividing the four test sessions (S1-S4) into eight sub-sessions Ss1-Ss8, this prediction was supported, with the first successful actions in all nine groups occurring at L1 and success at higher levels building over time, despite children knowing the most valued rewards to be available at higher levels from the start (figure 2). The percentage of available rewards retrieved was initially highest and almost exclusively at L1, with no significant change across the eight sub-sessions (L1 retrieval: Ss1 = 95%, Ss8 = 89%; Pages trend test, $L = 1510.5$, $k = 8$, $n = 9$, $p = 0.27$; figure 2). In contrast, success at L2 began to rise only later, showing a significant upward trend from a mean of 5% at Ss1 to 82% at Ss8 (Pages $L = 1594.5$, $k = 8$, $n = 9$, $p = 0.001$); success at L3 took off even later but then also rose significantly, from zero at Ss1 to 57% at Ss8 (Pages $L = 1575.5$, $k = 8$, $n = 9$, $p = 0.014$; figure 2). These ‘overview’ data suggest that learning about the affordances of SW was overall markedly cumulative.

(ii) Did cumulative learning occur at the group level? We distinguished three grades of cumulative learning (individual and/or social) that might occur in the SW environment. The lowest grade occurs where an exit success at one level is followed by and thus plausibly facilitates success on a different exit at the same level. The middle grade is where an exit success at one level is directly followed by success on a different exit at a higher level. This meets three criteria for cumulation: (i) experiences at the lower level,

such as discovering that, and how, the reward eggs can be manoeuvred and released can *facilitate* successful manoeuvres at the higher level; (ii) the higher level requires *additional* skills (fashioning and/or using of tools) and (iii) the higher level *delivers greater rewards*. Together these are common criteria for cumulative technological progress, whether culturally driven or not, and are the minimum criteria applied in the present study. Finally we recognise and distinguish a higher grade of cumulation in which the ascent of levels concerns the *same* kind of exit; here, the same three criteria listed above are met, but the first takes a stronger form because experiences of the exit concerned are clearly directly relevant to mastery of this exit at the higher level, given the additional challenges of tool use are met.

Accordingly, focusing initially on only the mid-level grade of cumulation, we note that all groups achieved their first successes at any level in the order L1 then L2 then L3 (two-tailed binomial test: $p < 0.004$), as predicted by the hypothesis that cumulative learning was at work, with earlier achievements able to facilitate more challenging ones at higher levels. Arrows in figure 3 represent sequential transitions between group-level first successes at the different exits (arrow heads) and the exit where the immediately prior success had been achieved. The hypothesis that cumulative learning was at work would predict that the flow represented in this way would be predominantly upwards through the levels, even though high value rewards were available at the highest levels from the start. This was confirmed. Overall, there were 33 upward transitions versus 5 down (two-tailed binomial test: $p < 0.001$). Such directional trends were significant both for the more specific transitions to L2 (23 up versus 1 down; two-tailed binomial test: $p < 0.001$), and to L3 (10 up versus 0 down; two-tailed binomial test: $p = 0.002$; figure 3a). Additional analyses that focus on the transitions between levels are provided in the supplementary information.

(iii) Did cumulative learning occur at the individual level? The pattern of cumulative progress revealed at the group level was also evident at the level of individual children, with all but one child achieving lower level success before higher (L2: 46 upward v 1 downward; two-tailed binomial test, $p < 0.001$; L3: 21 upward v 0 downward; two-tailed binomial test, $p < 0.001$). The direction of the specific transitions to each level illustrated in figure 3b were significantly positive (overall: 78 upward v 23 downward; two-tailed binomial test: $p < 0.001$; L2, 56 upward versus 5 downward; two-

tailed binomial test: p < 0.001; L3, 22 upward versus 0 downward; two-tailed binomial test: p < 0.001; figure 3b). Additional analyses that focus on the transitions between levels are provided in the supplementary information.

(b) Testing the prediction that groups will make greater progress than asocial and L4-only controls

(i) Groups versus Asocial Controls. The asocial control children tested individually were markedly less successful in their achievements, with none reaching L3, whereas 16% of children managed this in the groups. Similarly, 48% succeeded at L2 in groups, versus only 17% in the asocial condition. Statistical comparisons between the two conditions began from the constraint that the maximum interaction of any asocial child with the SW (i.e., total time an individual spent manipulating a reward egg, or exit) totalled 15 minutes in the asocial condition, even for those who experienced multiple sessions. We therefore computed the number of minutes each child spent interacting with the SW, in both the group and asocial conditions, prior to their first success at each level, L1-L3. We then compared the proportion of successful children in the two conditions after the first 1, 5, 10 and 15 minutes of interaction with the SW for each child involved (figure 4, see supplementary table S4 for descriptives).

At L1, a greater proportion of children gained rewards within the first minute of their task interaction when in groups rather than acting alone (Fisher's exact test, p = 0.015; figure 3). With greater interaction times success rates at L1 tended to converge (5 minutes: Fisher's exact test, p = 0.080; 10 minutes: Fisher's exact test, p = 0.746; 15 minutes: Fisher's exact test, p = 0.927; figure 4), so more time engaged with the SW appeared equally beneficial to L1 success in the two contexts.

In contrast, and of key relevance to the issues at the heart of our study, at L2 the proportion of group and asocial participants who succeeded on the task became ever more divergent, with the proportion of children in groups succeeding at L2 increasing much more rapidly than in asocial controls (where the proportion succeeding rose little), as interaction time increased (1 minute: group = 6%, asocial = 4%; Fisher's exact test, p = 1.000; 5 minutes: group = 26%, asocial = 9%; Fisher's exact test, p = 0.109; 10 minutes: group = 39%, asocial = 13%; Fishers exact test, p = 0.018; 15 minutes: group = 44%, asocial = 17%; Fisher's exact test, p =0.002; figure 4). At L3 only children in groups were successful, with the proportion successful increasing significantly across the 15 minute interaction period (1 minute: 2% successful; 15 minutes: 11% successful; Fisher's exact test, $p = 0.002$; figure 4). The children working individually did not succeed, or even attempt, L3, suggesting no significant benefit to increasing exposure.

(ii) Groups versus L4-only Controls. The L4-only control group (L4C) achieved no success at the only box-shelf available, L4, but 12 of the 15 children in the group attempted the task. Given that no children in either condition succeeded at L4 it is more instructive to examine the quality of *attempts* focused at L4 in the various groups, with attempts directed merely towards the reward deemed less advanced (less close to success) than exit interactions . The ratio of actions on exits versus merely manipulating the egg was as low as 4:109 in L4C, whereas the ratio was significantly higher (149:347) for the nine main groups (Fishers exact test, $p < 0.001$ two-tailed), a pattern of performance indicative of more advanced interactions in the groups.

These results show that the group context facilitates children's cumulative achievement, as their experiences at one level are built on at higher levels. However like the results in section (a), they do not yet directly address the further question of the role of social learning, which we now move to.

(c) Testing the prediction that social learning would be common

(i) Patterns of first exit use in groups. The majority of the nine groups (independently of group size; see supplementary table S5) achieved successful reward extraction from all four L1 exits, whereas by the time L3 was frequently attained there was a marked exit preference, with all groups successful on the trapdoor and just one group successful on the ramp ($\chi^2(1) = 6.400$; $p = 0.011$). Results at L2 were somewhat intermediate (figure 3a). This progressive convergence on the trapdoor, which these results suggest was the most effective to exploit, became most pronounced (number of individual first successes at trapdoor exit markedly higher in comparison to the other exits) when analysed at the level of individual children, as the first inventions in the group described above became shared innovations amongst others, implicating social learning (L1: $\chi^2(3) = 8.63$; $p = 0.035$; L2: $\chi^2(3) = 57.24$, $p < 0.001$; L3: $\chi^2(1) = 19.17$, $p < 0.001$; figure 3b).

(ii) Patterns of first exit use: group versus asocial conditions. The distribution of first inventions across exits, as expected, did not differ between the asocial and social

groups at either L1 (Fisher's exact test $p = .31$; Figure 3a,c), or L2 (Fisher's exact test $p = .73$; Figure 3a,c). However, when instead all first successes (when children could have watched others and adopted the options displayed) were compared, the difference between the asocial controls and the social groups was significant at L1 (Fisher's exact test, $p = .045$ one tailed; Figure 3b,c), although not at L2 ($p = .71$; Figure 3b,c), thus providing evidence that social learning influences were creating homogeneity of innovations in groups at L1, absent in the asocial condition. Additional analyses that show the relative difficulty of each exit (i.e. time to first success) are provided in the supplementary information.

(iii) Exit use observed prior to first success. Our principal means of inferring social learning in the open diffusion group context is through correlations between what children witness and the actions they subsequently take. First inventions of an exit success in a group were often observed by others (see supplementary information for coding details) , with an average of 2.55 other children observing each of these first solutions (range 0-7). Figure 5 illustrates the predominance of cases in which children had watched a prior success on a particular exit before they then achieved their own first success at that same exit. For several of these, the number of occurrences is sufficient to test the ratio statistically, revealing that a majority of children achieved their own first success at an exit where they had previously watched the success of another child (ramp: Fisher's exact test, $p < 0.001$; trapdoor: Fisher's exact test, $p < .001$; pushdoor: Fisher's exact test, $p = 0.008$; lift: Fisher's exact test, $p = 0.007$; figure 5). The link between observing, and subsequently matching, successful exit solutions was weaker at L2 and L3, reaching significance only for the trapdoor exit (Fisher's exact tests, L2, $p < 0.001$; L3, $p = 0.022$), but this reflected the fact that most efforts were directed at this exit, so provided the most instances other children could view. Additional analyses that consider all exits observed prior to first success are provided in supplementary table S6, with similar results.

Complementary analyses show that children's first success at the L1 exits was predicted by whichever exit use they had viewed *most* beforehand, providing strong evidence for social learning (figure 6: see supplementary information for statistical analysis). The effect was reduced at L2 and L3, given the increasing focus on one exit.

(d) Testing the prediction that progress would be cumulative at the group level (cumulative culture)

(i) Identity of inventor at first success. An important feature of cumulative culture is that what is achieved by some individuals in earlier phases can be built on in later phases by others, a criterion we next address. Figure 7 shows the identity of the child who in each group was the inventor of the first solution for each exit at each level. In some cases the same child was serially responsible, their progress through the levels thus providing circumstantial evidence for cumulative learning at this individual level, rather than at the group level, via social learning. However, in a majority of cases (31/41: binomial test, $p < 0.001$) the inventor was an individual different to the child who had been the inventor at the prior level. In conjunction with the data showing that most children replicated the exit success they observed the initial inventor achieve, the overall picture is thus one of cumulative cultural progress insofar as an invention at one level spreads to become a shared innovation, with a new inventor then building on it to create a further higher-level innovation, shared in turn by others in the group. Figure 7 includes cases where this process sometimes extended to L4, an effect found in the supplementary experiments we now describe in a Phase B of the study. All predictions and results of Phase A are summarised in supplementary table S1.

8. The facilitation of cumulative culture: ecological challenges and better innovators (Phase B)

The results reported above provide the first evidence consistent with the hypothesis that a limited capacity for cumulative culture is developing in young children, with the dual engines of invention and social learning permitting escalating achievements in a micro-society and micro-world (the SW). These results thus contrast with those of the only prior diffusion experiment testing the hypothesis [45]. However, in Phase A no child had attained the highest level of reward (L4) available. We therefore conducted supplementary experiments described next, in which we tested the potential effects of two elements that are considered important in shaping cumulative cultural evolution in the real world: (i) the occasional emergence of more gifted innovators (in this case instantiated by the inclusion of older children); and (ii) ecological change that escalates the selection pressures for cumulative culture to advance.

(a) Participants, materials and procedures

For the advanced models intervention, two nursery groups (KM and KA) that had already participated in the experiments described above, but reached only L3, continued through three further one-hour sessions, S5-S7, each of which contained a 5 minute rebait period (continuing the rebait regime employed in S4). In the first half of S5 and S6, two six-year-olds from the same school were introduced into the group. These older models were invited simply to 'work on the puzzle with the younger children', allowing the potential diffusion of any novel solutions they contributed to be as naturalistic as possible. The models were not present for the second half of both S5 and S6, and were also absent in S7, permitting other children to freely exploit any more advanced techniques they had seen the older children use. For further details see supplementary methods.

For the ecological change intervention, L1, L2 and L3 were removed, simulating an 'ecological catastrophe'. This was also done in the L4-only Control condition described, and analysed in comparison to our earlier Phase A, the results of which are now compared directly to those of our ecological change groups below. Additionally, in order to determine how far children could be pressed by the ecological change, new opportunities were opened up as an additional box-shelf (L5) designed to be yet more challenging than L4, was substituted in the L3 slot. To succeed at L5 it was necessary to swing a screen out of the way using one tool, to allow a hook tool to be used to catch a loop and open the exit, providing a further opportunity for cumulatively building on what had been learned earlier. In sum, children were presented with the new affordances of L4 and L5, with the easier pickings of L1-L3 now unavailable. Two nursery groups (PA and PM) earlier involved in S1-S4 but reaching only L3 were presented with the SW for two 1-hour sessions, S5 and S6, with success or failure on the task determining the structure of a further two sessions. If a group successfully released a reward at either L4 or L5 during sessions 5 or 6 the group was allowed a further two 1-hour sessions, S7 and S8, providing an opportunity for solutions to diffuse in the group. If instead a group failed to release a reward during sessions 5 and 6, two older (6-year-old) children were to be introduced into the group for the first 30 minutes of session 7, before the younger children were allowed to interact with the task on their own for the remainder of session

7, and session 8. A 5-minute rebait period was employed throughout S5-S8. For further details see supplementary methods.

9. The facilitation of cumulative culture: Phase B Results

(a) Effects of advanced models

The older child models succeeded in extracting rewards not only at L1-L3 but at L4, providing the first demonstrations of how to solve L4 to the two nursery groups (see supplementary table S7 for descriptives). Moreover three of the younger children in group KM then went on to succeed at L4, adopting the method used by the older children, which required extending the unfold tool, including its terminal hook element, and using this to catch loops at L4. No child had been seen to do this in the previous eight sessions, indicating they copied from the older child. However, the first of the younger children applied the hook successfully to their favoured exit, the trapdoor (figure 8), not the ramp that had been opened by the older child, and the other two children then did so too. Accordingly they displayed a cumulatively built technique that then acquired the status of a shared innovation. The older child's invention also impacted the group in ways not restricted to success at L4, with a total of 4 children attempting to use a hook, successfully or not (N hook uses prior to first success = 0, N hook uses after first success = 47; (two-tailed binomial test: $p < .001$), on ramp, trapdoor and lift; with 7 children increasing their level of solution from that witnessed prior to the introduction of the older models (four children advanced from L2 to L3, two children from L3 to L4, and one child from L1 to L4). By contrast children in the KA group failed to learn from the older child's hook use, but this modelling had taken place later (first L4 success session S6) than in group KM (first L4 success session S5), and as a consequence fewer successes occurred in total in KA than in KM (2 versus 9 successes), providing more limited opportunity for successful hook use to be observed. Overall these results underline the challenging nature of L4 for children of the ages common in our nursery groups.

(b) Effects of major ecological challenge

With L1-L3 unavailable, one child in group PA succeeded at L4 by inventing an unfold hook tool as described in the previous section (see supplementary table S8 for descriptives). However in this case, despite an average of 2.07 children watching each of this child's successes, none adopted the technique. There were also successes at the L5 trapdoor through one child applying a long combine tool and another child (the same individual who had previously solved L4) later using an unfold tool with hook, but these also did not spread to become shared innovations in the group, despite the first success being watched by nine children and successes being viewed on average by 2.67 children.

In contrast no success occurred at L4 or L5 in two hours of exposure in group PM. Accordingly the plan outlined above was activated to introduce two older models to this group, one of whom extended the unfold hook tool and used it to open the L4 ramp exit. One younger child then repeated this to open the ramp exit, but also went on to apply this tool to open the trapdoor, after which a further child did so (figure 8). Thus as in the advanced models intervention with L1-L4 in place described in section (a) above, these children employed a cumulatively built technique that later acquired the status of a shared innovation. Attempting the challenging use of this tool, even if without success, spread amongst a further seven children on ramp and trapdoor, but not the other exits (N hook uses prior to first success = 2, N hook uses after first success = 41; (two-tailed binomial test: $p < .001$) (figure 8; and see supplementary figure S2 for an overview of these results).

Given these results it is not surprising that the L4-only Control group (L4C) had achieved no success at L4 (Section 7(b) above). However, despite the lack of success it is instructive to examine the quality of *attempts* focused at L4 in the same way as in Phase A, but here focusing only on PA and PM. The ratio of actions on exits versus manipulating the egg was as low as 4:109 in L4C, whereas in PA and PM it was much higher, at 33:36 and 59:56 respectively. Combining the latter two similar results, the ratio was significantly higher (Fishers exact test, $p < 0.001$ two-tailed) in these groups that had experienced the whole hierarchically structured SW in the earlier sessions S1-S4, and provided evidence of cumulative cultural progress in that period. However, equivalent ratios expressed at the level of individual children fail to achieve significance: L4C, 2:12; PA/PM 13:19; $p = 0.076$. This makes the interpretation of these contrasts complicated yet instructive, and we return to them in the discussion below.

Results of Phase B, (a) and (b) are summarised in supplementary table S2 using the same framework distinguishing predictions and associated results as in Phase A.

10. Discussion

Introducing a set of novel experiments in the field of cumulative culture, we have described a cluster of the findings most relevant to the present theme issue of this journal. As we noted in our introductory review, the literature addressing this topic through experiments remains limited, especially with children. The latter have focused on cumulative social learning, but either omitted the wider cultural diffusion necessary to identify 'innovation' as here defined [26,46,50], or employed a task that necessarily required cumulative steps for success, constraining inferences about the importance of cumulation in progressive achievements [47], or offered a task in which children could obtain no rewards other than transferring maximal quantities of rice between two containers, as an experimenter encouraged them to attempt [48]. Children showed little or no evidence of cumulative culture in the latter experiment; however, most could already spontaneously recognise adequate ways of transferring rice from the start.

Accordingly we suggest that through our experiment we have obtained some of the first experimental evidence that (limited) capacities for cumulative cultural progress are already developing in young children. This necessarily required the dual elements of invention and social learning, leading to shared innovations, some of which were cumulatively built as progress was made in achievements at higher and more challenging (and more rewarding) levels in our Small World environment. However, our study also emphasises the challenges of drawing inferences about these processes within the interpretive complexities of cumulative cultural experiments. We discuss these before focussing on particular implications of our results.

(a) Defining and identifying invention, innovation, social learning and cumulative cultural change: challenges for cultural diffusion experiments

The emergence of shared innovations requires the prior *emergence of a novel element*, which is how we have here defined an *invention*. For objective clarity we treated each first success at opening a specific exit among the 16 available in the SW as such an invention, and at successive levels in the SW these required the invention of different kinds of tool use and novelties in how to apply these to exits familiar from other levels (as in fashioning longer tools for L3 and hook tools for L4). We think this a satisfactory approach insofar as the term 'invention' is thus technically and clearly defined in the context of our study, and the present journal issue [11]. However, calling actions like using a finger to press down the trapdoor at L1 'invention' may seem to over-dignify it in comparison to the word's everyday meaning and cultural instances like the invention of light bulbs and cyclotrons. What we may call 'inventions' (or innovations) vary in the size of the 'leap' from what existed before, to what exists after, the invention, and accordingly perhaps the cognitive capacity required [56,57]. Such contrasts surely exist within our study: for example, compare the L1 finger-push on the trapdoor with the inventive fashioning of a hook tool and its use to catch a loop on the ramp door and pull it down at L4. For the moment we simply acknowledge this issue, but perhaps in future, progress may be made in distinguishing levels or kinds of inventions, the nature of the cognitive gap between them and the nature of the mental leap needed to bridge the gap.

Identifying social learning in the present experimental context is also not straightforward. Robust experiments can be designed to test for social learning, by seeding alternative actions in models viewed by others in dyadic or diffusion designs, and/or by adding asocial controls [26-28]. But in such instances, the experimenters contribute the 'inventions', whereas to test for cumulative culture in the round, we need to allow the participants themselves to contribute inventions as well as social learning. Inevitably this removes an element of experimental control and necessitates reliance on more correlational or circumstantial evidence about whether social learning picks up on the invention to transform it into a shared innovation. These interpretive challenges are further compounded when interest extends to cycles of invention, social learning and innovation, of the kind essential to extended cumulative culture.

With these inherent challenges in mind, we discuss the inferences we draw from our study, and future elaborations and refinements of the approach

(b) Innovation and cumulative cultural progress in a child micro-society

The most compelling conclusions about the principal processes under investigation come through multiple, but convergent, lines of evidence within our results. The fundamental characteristics agreed upon by most commentators on cumulative culture are that group achievements that are shaped through social learning are later built on by others to create innovations that are then further spread through social learning [2, 3, 13-15, 32, 42]. Evidence for these effects came from contrasts between our asocial and group conditions, and the power of what children saw others do to predict what they elected to do themselves (Phase A (sections c,d) and Phase B (sections a,b); figures 5-8), which we discuss further below. A further criterion emphasised by some is that later stages come to be beyond what any individual can develop from scratch [3, 13], confirmed in our level-4 only condition (Phase A (section b)).

Despite ceding the control of inventions to the children rather than seeding them experimentally ourselves, there was clear evidence of social learning already in the extent to which children's success at particular exits among the four at L1 was predicted by their prior observation of the focus of other children's actions (figures 6 and S3). More generally, across levels L1-L2-L3 children were much more likely to match an earlier witnessed first invention of a successful solution in their group, than invent an approach themselves (figure 5). Figure 5 illustrates that overall, 81% of children's first successes were those first invented by another child in their group. This figure is remarkably similar to the 84% of children being classed as 'followers' (social learners) as opposed to 'innovators' in one of the few other child micro-society cultural diffusion experiments [31] and a similar figure of 88% for 'followers' in a recent study with multiple solution options available [22] .

Additional evidence that the *cumulative* learning indicated in figures 1 and 4 included elements of cultural, as well as individual, learning derive from other convergent results. The first innovation at any exit always followed an ascending order in the group, L1-L2-L3 (including rare cases of L1-L3), and for all the 9 principal exits used, the majority of children achieved their own first success after witnessing that first success in their group (wherever frequencies were adequate for testing, they displayed a statistically significant effect). The typical occurrence at L2 was for one child to achieve a first success in their group having already gained experience on their exit of choice (and often other exits) at L1, but with that experience having followed observation of a *different* inventor at L1 (figure 7). The new approach at L2 was then copied by other children, to become a shared innovation. This was repeated at L3 but almost exclusively on the trapdoor exit (figures 7 and 8). Such effects went significantly beyond those seen in the asocial condition, demonstrating the power of cycles of invention, social learning and creation of escalating levels of innovation within social groups, already manifested in this early childhood micro-society. Such effects echo those demonstrated in cultural transmission experiments involving adults faced with complex technological challenges, in which communities (of five individuals) achieved significantly greater overall cumulative success than when each successive generation involved only a single individual [42; 39-41, for related findings]. Henrich and Muthukrishna have referred to the cognitive resources shared in the group condition as those of a 'collective brain' [3, 42], an expression apt to the contrast between the progressive innovations of the children in our groups, compared to the asocial condition. We suggest there is much scope here for linkage with recently expanding research on collective cognition in animal groups [59, 60].

The results of the second series of experiments that either introduced older children capable of higher-level invention, or imposed ecological change that removed the hitherto easiest options of L1-L3, are particularly relevant for such contexts. In group KM, after an older child fashioned a hook tool and used it to open the L4 ramp exit, another child copied the application of this tool but used it on the most popular exit for their group, the trapdoor, and this was adopted by two further children, thus creating a shared innovation. A similar sequence occurred in group PM, where success at L4 occurred after the group was deprived of L1-L3, but only when an older child was again introduced. One younger child copied that child's application of a hook tool to the ramp, but then, as in KM group, used it on the popular trapdoor exit, which was in turn replicated by a further child, again creating a shared innovation. In these episodes we begin to see the key components of cumulative culture, involving cycles of invention, social learning and shared innovations, with the inventions often being made by different individuals at escalating levels of achievement, thus building on the prior progress of the group as a whole.

We note that in group PA the invention of hook use by a young child was not copied by others, unlike when the model was a six-year-old in group PM and was copied. This is consistent with a recent systematic demonstration of a preference in young children for copying older as opposed to younger models [61].

(c) Omissions and future research

The experimental approach introduced here has generated a rich dataset. In this article we have chosen to focus on the aspects of the results that best relate to the innovation theme of the journal issue, leaving aside several other significant questions that can be addressed. One major example of the latter concerns the emergence in the different child nursery groups of cultural variants. That such effects occurred is indicated here in figure 7, where it is evident that some groups omitted any focus on certain exits on which other groups were active; however, the more comprehensive treatment that this question merits is beyond the scope of the present article and will be addressed elsewhere. The same is true for our comparative studies of how other species behave in SW and analogous environments.

Earlier we explained why for this exploratory study we chose an open diffusion design. Doing so has generated a wealth of data that begin to chart the scope and limits of cumulative cultural and individual learning in children operating in a complex environment, but many of our inferences about the crucial social learning elements must remain tentative because they are essentially based on correlations between what children see and what they subsequently do, rather than on experimental manipulations. The social effects indicated by our results may have ranged over the potentially motivating effects of merely witnessing other children engaging with the SW, including particular parts of it, and also by seeing others gain rewards, although we have offered evidence for more specific copying of exit usage and application of tool use in doing this. We suggest that what we have learned in the present study can now best be built on by applying other forms of diffusion experiment that can pin down cumulative cultural transmission specifically, with more rigour. For example, if the SW were run in a replacement design [26, 27, 32, 34] with triplets of children in each generation, then starting with children A, B and C, three successive replacement rounds would lead to a community made up of different children, D, E and F. If on average the later triplets displayed higher levels of achievement than the initial ones, it would unambiguously demonstrate cumulative cultural progress. We hope the present study will provide significant foundations for the further research that is needed to understand better the developmental origins of innovation and cumulative culture.

Data accessibility. Data file is available in the supplementary information.
Author's contributions. AW conceived the study. AW, NM, EB, LD, AL, GL collaborated in detailed experimental design. EB, LD, AL, VB executed the study and coded videotapes. NM and AW conducted statistical analyses. NM and AW drafted the manuscript, then co-authored and approved by all.
Competing interests. The authors declare no competing interests.
Funding. This research was funded by grant ID40128 from the John Templeton Foundation to AW and K. Laland.
Ethics. All experiments were approved by the University of St Andrews ethics committee, UTREC.
Acknowledgements. We thank all the parents who gave their ethical consent for these whole-group studies, and the teachers and children involved. Thanks to Keith Haynes and his team for construction of apparatus and to Keelin Murray for assistance in video coding.

References

1. Whiten, A., Hinde, RA, Stringer CB, Laland KN 2011 Culture Evolves. *Phil. Trans. R. Soc. B* **366,** 938-948.
2. Richerson PJ, Christiansen MH (eds). 2013 *Cultural evolution: society, technology, language and religion.* Cambridge MA, MIT Press.
3. Henrich J. 2015 *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter.* Princeton NJ, Princeton University Press.
4. Whiten A, Ayala FJ, Feldman MW, Laland KN. 2017 The extension of biology through culture. *Proc. Natl. Acad. USA* in progress.
5. Aplin LM, Farine DR, Morand-Ferron J, Cockburn A, Thornton A, Sheldon BC. 2015 Experimentally induced innovations lead to persistent culture via conformity in wild birds. *Nature* **518,** 538-541
6. Whitehead H, Rendell L 2015 *The Cultural Lives of Whales and Dolphins.* Chicago, Chicago University Press.
7. Alem S, Perry CJ, Zhu X, Loukola OJ, Ingraham T, Søvik E, Chittka, L. 2016 Associative mechanisms allow for social learning and cultural transmission of string pulling in an insect. *PLOS Biol.* **14,** e1002564.
8. Dukas R 2017 Cognitive innovations and the evolutionary biology of expertise. *Phil. Trans. R. Soc. B*
9. Reader SM, Laland KN. 2003 *Animal innovation.* Cambridge, Cambridge University Press.
10. Reader SM, Morand-Ferron J, Flynn EG 2016 Animal and human innovaton: novel problems and and novel solutions. *Phil. Trans. R. Soc. B* **371,** 20150182.
11. Hochberg M, Marquet P, Boyd R, Wagner A. 2017 Innovation: an emerging focus from cells to societies *Phil. Trans. R. Soc. B,* introductory paper to this issue.
12. Legare CH, Nielsen M 2015 Imitation and innovation: the dual engines of cultural learning. *Tr. Cog. Sci.* **19,** 688-699.
13. Tennie C, Call J, Tomasello M. 2009 Ratcheting up the ratchet: on the evolution of cumulative culture. *Phil Trans R Soc B* **364,** 2405-2415.
14. Pagel M. 2012 *Wired for culture: Origins of the human social mind*. New York, NY, WW Norton & Company.
15. Dean LG, Vale GL, Laland KN, Flynn E, Kendal RL. 2014 Human cumulative culture: a comparative perspective. *Biol Rev* **89,** 284-301.
16. Ridley M. 2015 *The evolution of everything: How new ideas emerge.* London, Fourth Estate.
17. Currie TE, Mace R. 2011 Mode and tempo in the evolution of socio-political organization: Reconciling 'Darwinian' and 'Spencerian' evolutionary approaches in anthropology. *Phil. Trans. R. Soc. B* **366,** 1108-1117.
18. Fogarty L, Creanza N 2017 The niche construction of cultural complexity: interactions between innovations, population size, and the environment. *Phil. Trans. R. Soc. B*
19. Pitt-Rivers A Lane-Fox. 1906 *The evolution of culture and other essays*. Oxford, Clarendon Press.
20. Basalla G. 1988 *The evolution of technology*. Cambridge, Cambridge University Press.
21. Rogers EM 2003 *Diffusion of innovations,* fifth edition. London, Free Press.
22. Carr K, Kendal RL, Flynn EG 2015 Imitate or innovate? Children's innovation is influenced by the efficacy of observed behaviour. *Cognition* **142,** 322-332.
23. Bruner JS, Jolly A, Sylva K. 1976 *Play – its role in development and evolution.* New York: Penguin.
24. Bateson P, Martin P 2013 *Play, playfulness, creativity and innovation.* Cambridge: Cambridge University Press.
25. Hopitt W 2017 The conceptual foundations of Network-Based Diffusion Analysis: choosing

networks and interpreting results. *Phil. Trans. R. Soc. B*
26. Mesoudi A, Whiten A. 2008 The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Phil. Trans. R. Soc. B*, **363,** 3489-3501.
27. Whiten A, Mesoudi A. 2008 Establishing an experimental science of culture: animal social diffusion experiments. *Phil Trans R Soc B* **363,** 3477–88.
28. Whiten A, Caldwell CA, Mesoudi A. 2016 Cultural diffusion in humans and other animals. *Current Opinion in Psychology* **8,** 15-21.
28. Horner V, Whiten A, Flynn E, de Waal FBM. 2006 Faithful replication of foraging techniques along cultural transmission chains by chimpanzees and children. *Proc. Natl. Acad. USA* **103,** 13878-13883.
30. Flynn E, Whiten A. 2008 Cultural transmission of tool-use in young children: A diffusion chain study. *Soc. Dev.* **17,** 699-718.
31. Whiten A, Flynn EG. 2010. The transmission and evolution of experimental 'microcultures' in groups of young children. *Dev. Psych.* **46,** 1694-1709.
32. Caldwell CA, Millen AE. 2008a Studying cumulative cultural evolution in the laboratory. *Phil. Trans. R. Soc. B* **363,** 3529-3539.
33. Marshall-Pescini S, Whiten A. 2008 Chimpanzees (*Pan troglodytes*) and the question of cumulative culture: an experimental approach. *Anim. Cogn.* **11,** 449–56 (2008).
34. Caldwell CA, Millen AE. 2008b Experimental models for testing hypotheses about cumulative cultural evolution. *Evol. Hum. Behav.* **29,** 165-171.
35. Caldwell CA, Millen AE. 2009 Social Learning Mechanisms and Cumulative Cultural Evolution: Is Imitation Necessary? *Psychological Science* **20**, 1478-1483.
36. Morgan TJH, Uomini NT, Rendell LE, Chouinard-Thuly, … Whiten A, Laland KN. 2015 Evidence for the co-evolution of hominin tool-making, teaching and language. *Nat. Communic.* 6, 6029 DOI: 10.1038/ncomms7029.
37. Zwirner E, Thornton A. 2015 Cognitive requirements of cumulative culture: teaching is useful but not essential. *Sci. Rep.* **5,** 16781. (doi:10.1038/srep16781).
38. Caldwell CA, Millen AE 2010 Conservatism in laboratory microsocieties: unpredictable payoffs accentuate group-specific traditions. *Evol. Hum. Behav.* **312,** 123-130.
39. Caldwell CA, Millen AE. 2010 Human cumulative culture in the laboratory: Effects of (micro) population size. *Learn. Behav.* **38,** 310-318.
40. Derex M., Beugin M-P, Godelle B, Raymond M. 2013 Experimental evidence for the influence of group size on cultural complexity. *Nature* 503 389-+. (doi:10.1038/nature12774).
41. Kempe M, Mesoudi A.2014 An experimental demonstration of the effect of group size on cultural accumulation. *Evol .Hum. Behav.* **35,** 285–290.
42. Muthukrishna M, Shulman BW, Vasilescu V, Henrich J. 2014 Sociality influences cultural complexity. *Proc. R. Soc. B* **281**, 20132511.
43. Kirby S, Cornish H, Smith K 2008 Cumulative cultural evolution in the laboratory: An experimental approach to the evolution of structure in human language. *Proc. Natl. Acad. USA* **105,** 10681-10686.
44. Claidière N, Smith K, Kirby S, Fagot J. 2014 Cultural evolution of systematically structured behaviour in a non-human primate. *Proc.R. Soc. B* **281,** 1797. (doi:10.1098/rspb.2014.1541).
45. Flynn E. 2008. Investigating children as cultural magnets: Do young children transmit redundant information along diffusion chains? *Philosophical Transactions of the Royal Society B*, **363,** 3541–3551. (doi:10.1098/rstb. 2008.0136).
46. Whiten A, McGuigan N, Marshall-Pescini S, Hopper LM. 2009 Emulation, imitation, overimitation and the scope of culture for child and chimpanzee. *Phil. Trans. R. Soc.* **364,** 2417-2428.

47. Dean LG, Kendal RL, Schapiro SJ, Thierry B, Laland KN. 2012 Identification of the social and cognitive processes underlying human cumulative culture. *Science* **335,** 1114–1148.
48. Tennie C, Walter V, Gampe A, Carpenter M, Tomasello M. 2014 Limitations to the cultural ratchet effect in young children. *J. Exp. Child Psychol.* **126,** 152–160.
49. Lyons DE, Young AG, Keil FC. 2007 The hidden structure of overimitation. *Proc. Nat. Acad. Sci.* USA **104,** 19751-19756
50. Reindl E, Apperly IA. Beck SR, Tennie C. 2017 Young children copy cumulative technological design in the absence of action information. Sci. Rep. **7,** 1788 DOI:10.1038/s41598-017-01715-2.
51. Beck SR, Apperly IA, Chapell J, Guthrie C, Cutting N. 2011 Making tools isn't child's play. *Cognition* **119,** 301-306.
52. Whiten A, Horner V, de Waal FBM. 2005 Conformity to cultural norms of tool use in chimpanzees. Nature **437**, 737-740.
53. McGuigan N, Whiten A, Flynn E, Horner V. 2007 Imitation of causally-opaque versus causally-transparent tool use by 3- and 5-year-old children. *Cog. Dev.* **22,** 353-364.
54. Mesoudi A, Whiten A, Laland KN. 2004 Is human cultural evolution Darwinian? Evidence reviewed from the perspective of 'The Origin of Species'. *Evolution* **58,** 1-11.
55. Köhler W. 2015 *The mentality of apes.* London, Routledge and Kegan Paul.
56. Cutting N, Apperly IA, Beck SR. 2011 Why do children lack the flexibility to innovate tools? *J. Exp. Child Psychol.***109,** 497-511.
57. Cutting N, Apperly IA, Chapell J, Beck SR. 2014 Why can't children piece their knowledge together? The puzzling difficulty of tool innovation. *J. Exp. Child Psychol.***125,** 110-117.
58. Nielsen M. 2013 Young children's imitative and innovative behaviour on the floating object task. *Infant Child Dev.* **22,** 44-52.
59. Couzin ID. 2009 Collective cognition in animal groups. *Trends. Cogn. Sci.* **13,** 36-43.
60. Biro D, Sasaki T, Portugal SJ. 2016 Bringing a time-depth perspective to collective animal behaviour. *Trends Ecol. Evol.* **31,** 550-562.
61. McGuigan N, Burgess V. 2017 Is the tendency to conform influenced by the size of the majority? *J. Exp. Child Psychol.* **157**, 149-165.

Figure captions

Figure 1. Apparatus. (a-d) The ‘Small World’. Rewards of escalating value inside plastic eggs could be extracted at potential exits on four increasingly challenging task levels. At each level, each of four exits required a different operation to release the egg, which then rolled out on the inclined lower plane. At Level 1, eggs could be pushed from a central position (a) to an exit by inserting a finger in a horizontal slit in the plastic sheet spanning the front of the SW, and exits operated to release the egg using a finger or simple stick tool (b-c). For the Ramp, the egg had to be pushed up a slope to the top of the ramp, then a trapdoor released below; for the Trapdoor, its front lip needed to be depressed; for the Pushdoor, the egg had to be pushed against a flap door at the back of the box-shelf; and in the Lift, the egg required to be pushed into a small box, that could then be lifted from the top so the egg fell out through a hole at the front (b,c). At Levels 2 and 3 the exits were similar, but Level 2 was 10 cm. away, so a stick tool had to be used to move eggs into position and operate the exits, and Level 3 was 30 cm away, so a longer tool needed to be created to open exits (d). At Level 4, hook tools had to be created to catch loops that operated the various exits. (e) Types of functional tool materials referred to in this paper. Top, green ‘Unfold’ types of tools; Below, yellow ‘Combine’ family of tools.

Figure 2. Percentage of available rewards retrieved at each level in successive sessions. Four one-hour sessions are here divided into eight half-hour sessions. The number of rewards available in each session was determined by dividing the total duration of each session (i.e., 30 minutes) by the rebait interval employed within that session (1-5 minutes). Pages tests reveal significant ascending trends for the more challenging levels 2 and 3 (see text for details).

Figure 3. Numbers of first successes at each exit and transitions between them. (a) First successes at group level. Arrows connect exits where a first group success occurred (arrow head) with the exit where the previous success was available to be witnessed in that group. Numbers inside boxes record the number of first group-level successes occurring there. (b) Individual child-level successes. Arrows connect exits where a first success occurred (arrow head) for an individual with the exit where the previous success was available to be witnessed by that child. Numbers inside boxes record the number of first individual-level successes occurring there. (c) First successes in asocial control children. Conventions as for (b). Note that only

16children contributed to these asocial condition results, whereas data from 101 children across 9 groups are represented in (b).

Figure 4. Percentage of children in group (G) and asocial conditions (A) achieving success at each level, as a function of minutes of their interactions with the SW.

Figure 5. Children's first successful extraction of a reward for specific exits, distinguishing cases where children had observed the invention of this in their group previously, from cases where they were the inventor. White bars represent the number of later first successes where the child had seen use of the precise exit; hatched bars represent first successes where the child had seen use of the exit at another level; black bars represent first successes discovered by the individual.

Figure 6. Exits used by children in their first successes as a function of their previous observation of others' successes at these exits. (a) L1; (b) L2; (c) L3.

Figure 7. Identity of inventor to achieve first success at each exit at each level. Wide black arrows indicate cases where the same child filled this role at successive levels (n =10). Wide white arrows mark cases where the inventor at a higher level was different to the child at the lower level for that exit (n = 31). Small arrows show diffusion to other children to create innovation at L4 (for other levels refer to figure 6). See text for discussion.

Figure 8. Exit and tool use before and after first invention of L4 solution in four groups. The mid-point shows the first L4 success (O = older child model; G = existing group member). Exits: red = ramp; blue = trapdoor; green = lift. Tools: yellow = combine tool; dark green = unfold tool, no hook; light green, unfold tool with effective hook added. Numbers distinguish responses, sometimes multiple, by different children. Numbers in bold indicate a first success.

Figure 1.

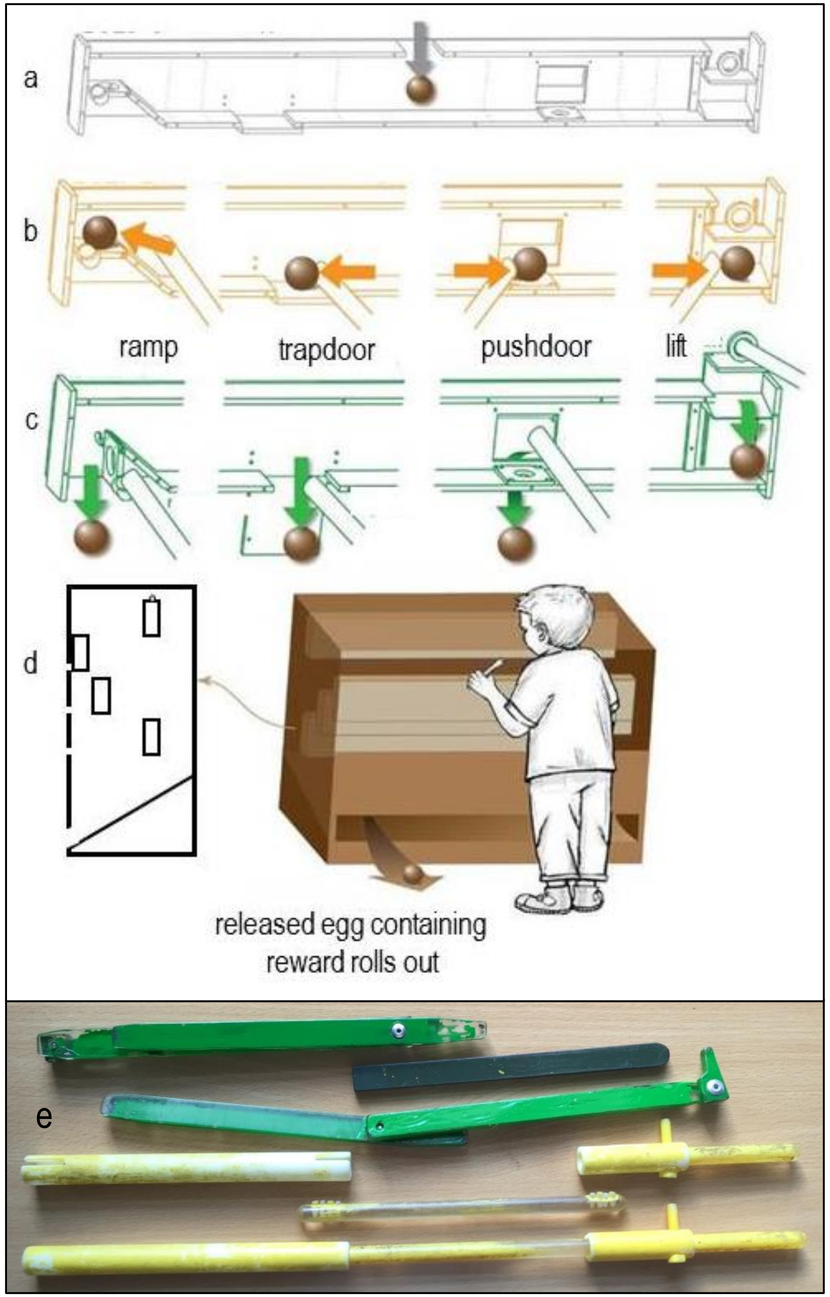
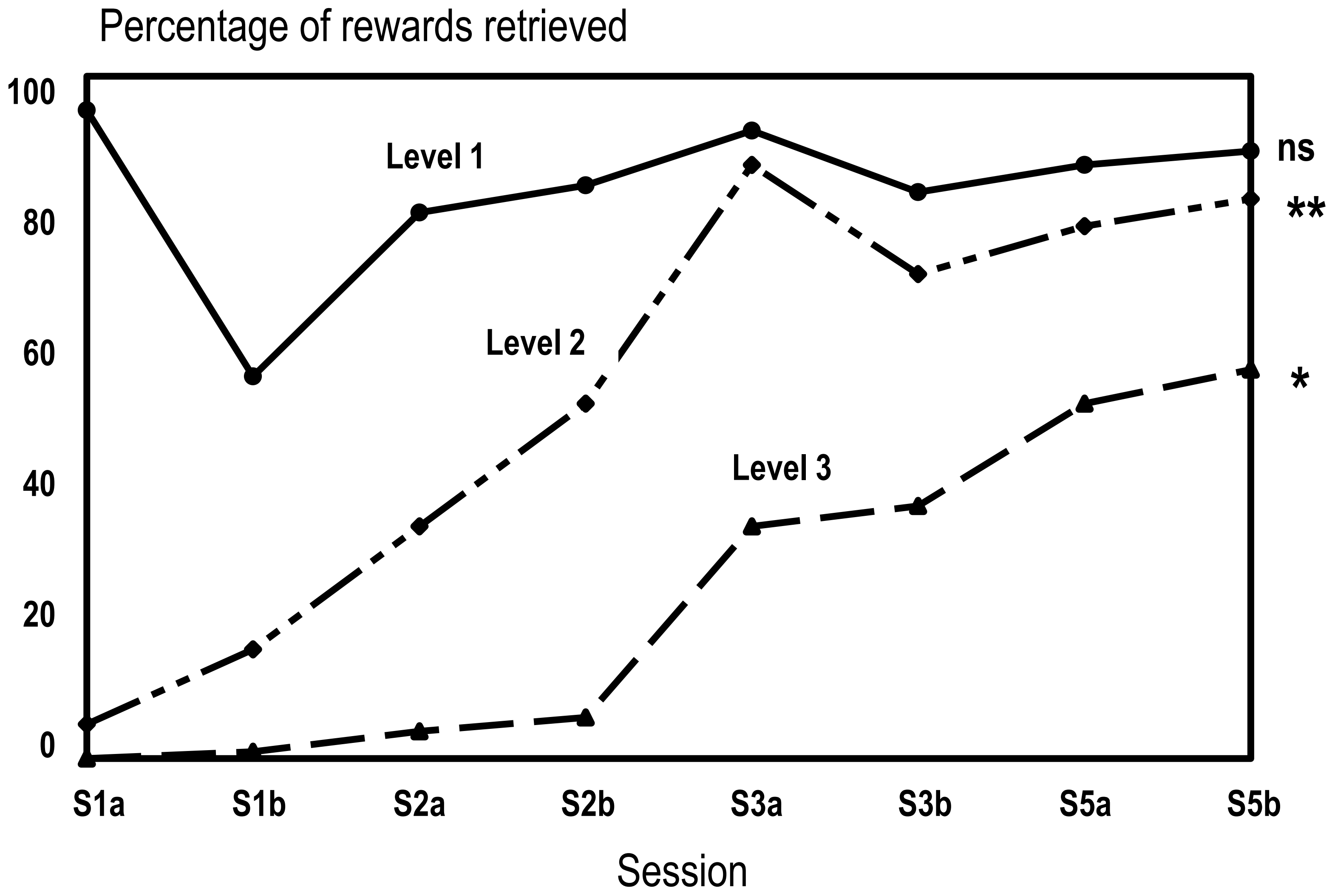
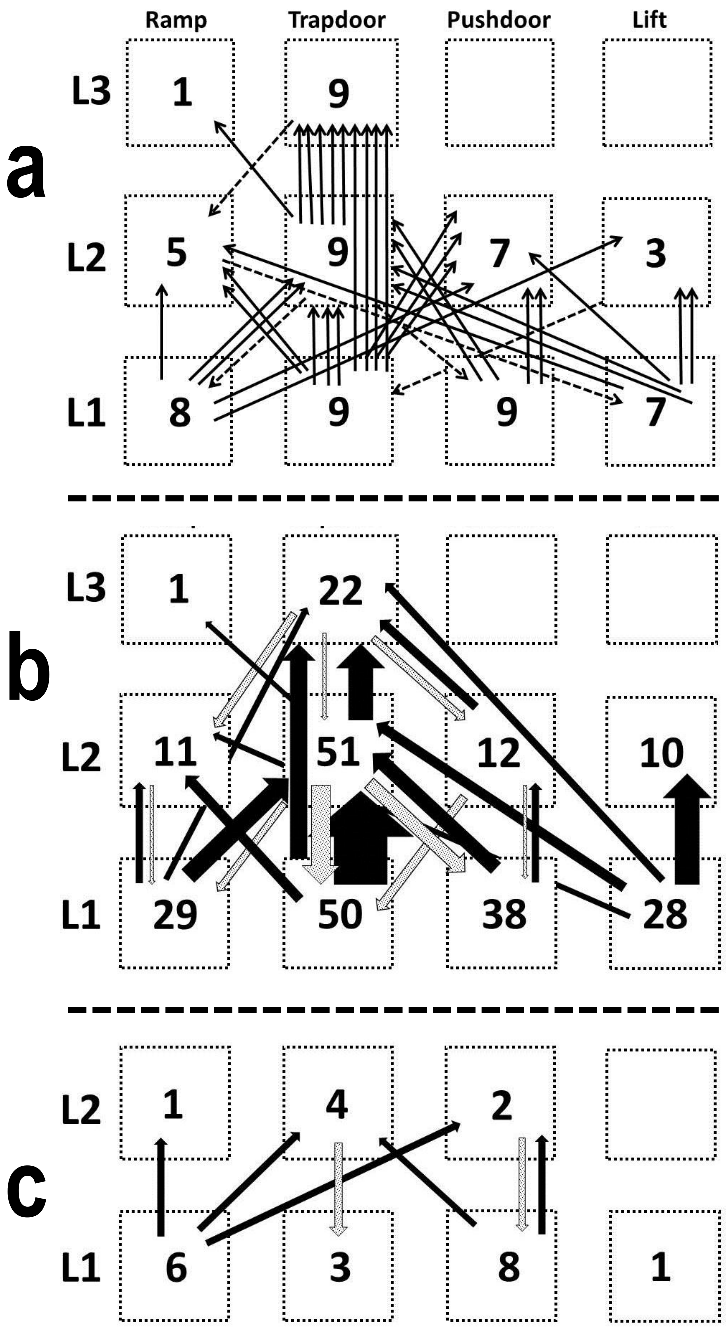

a
b
ramp
trapdoor
pushdoor
lift
c
d
released egg containing
reward rolls out
e

Figure 2.



Percentage of rewards retrieved
100
80
60
40
20
0
Level 1
Level 2
Level 3
ns
**
*
S1a
S1b
S2a
S2b
S3a
S3b
S5a
S5b
Session

Figure 3.



Ramp
Trapdoor
Pushdoor
Lift
a
L3
1
9
L2
5
9
7
3
L1
8
9
9
7
b
L3
1
22
L2
11
51
12
10
L1
29
50
38
28
c
L2
1
4
2
L1
6
3
8
1

Figure 4.

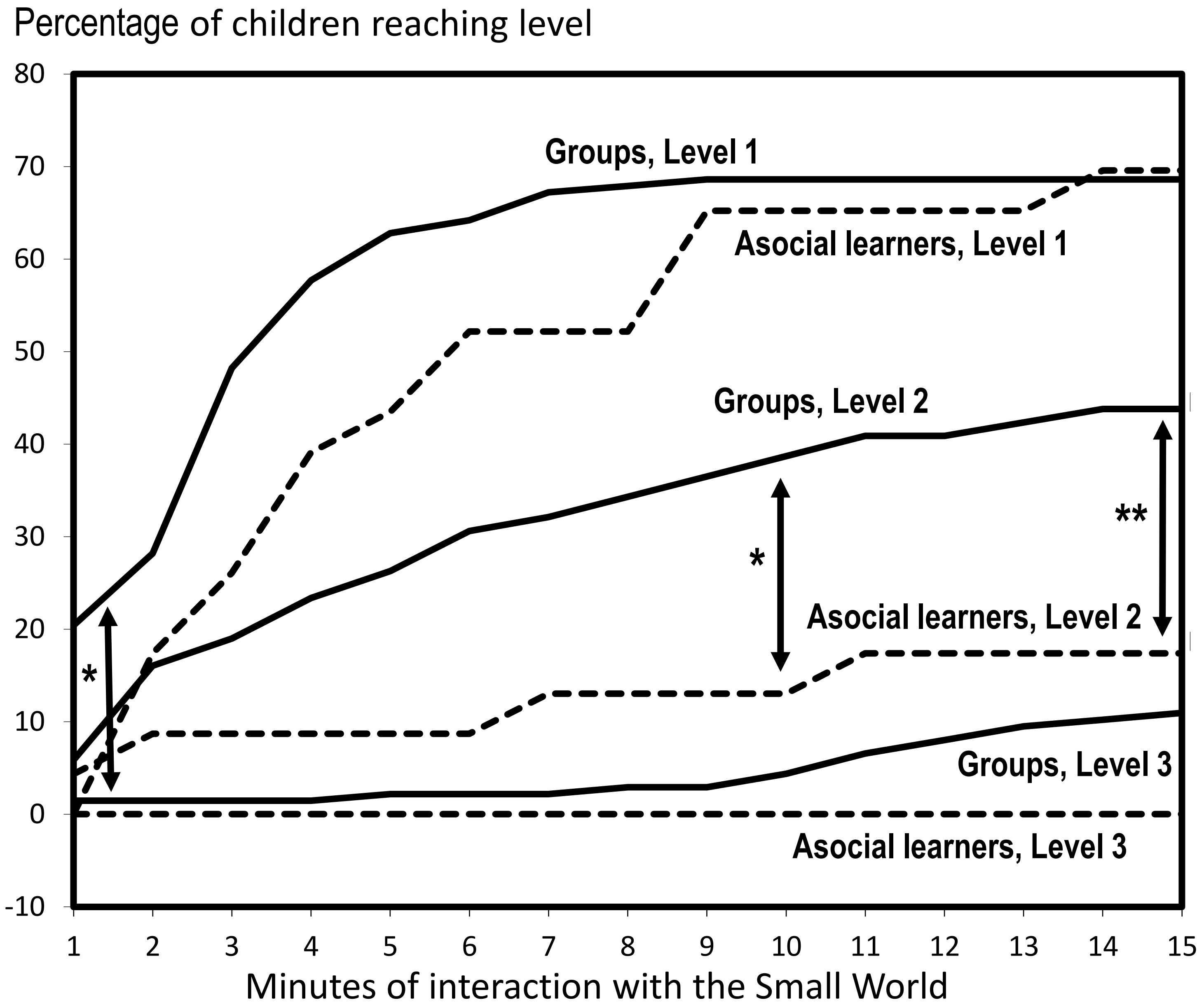
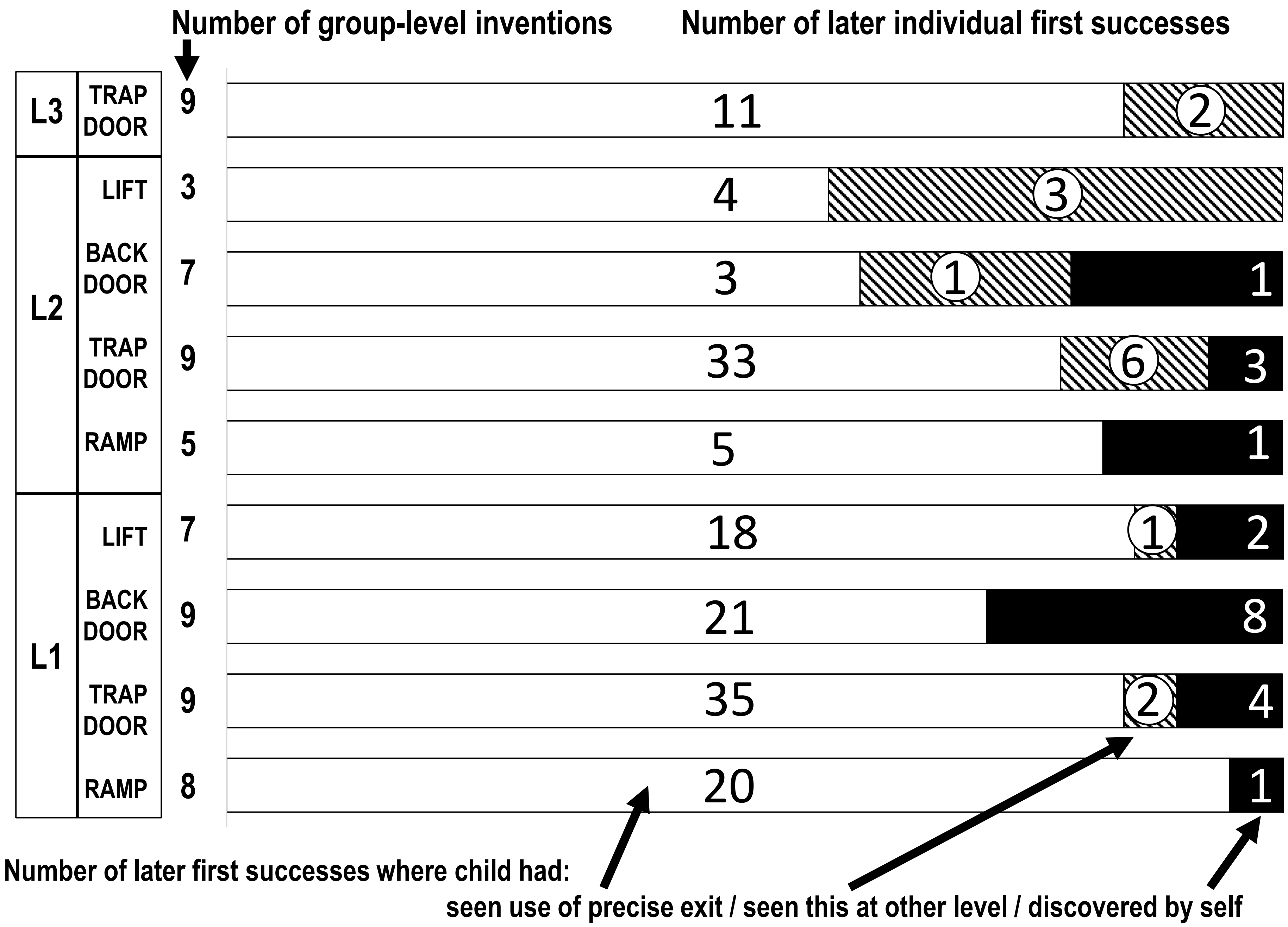

Percentage of children reaching level
80
70
60
50
40
30
20
10
0
-10
Groups, Level 1
Asocial learners, Level 1
Groups, Level 2
Asocial learners, Level 2
Groups, Level 3
Asocial learners, Level 3
*
*
**
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Minutes of interaction with the Small World

Figure 5.



Number of group-level inventions
Number of later individual first successes
L3
TRAP DOOR
9
11
2
L2
LIFT
3
4
3
BACK DOOR
7
3
1
1
TRAP DOOR
9
33
6
3
RAMP
5
5
1
L1
LIFT
7
18
1
2
BACK DOOR
9
21
8
TRAP DOOR
9
35
2
4
RAMP
8
20
1
Number of later first successes where child had:
seen use of precise exit / seen this at other level / discovered by self

Figure 6.

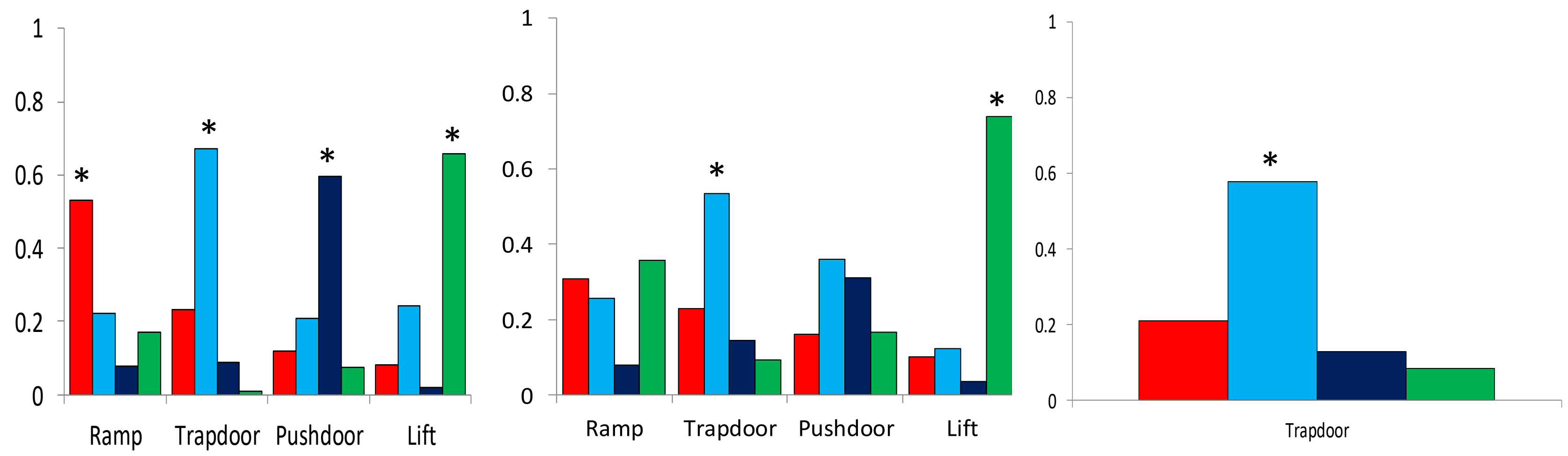
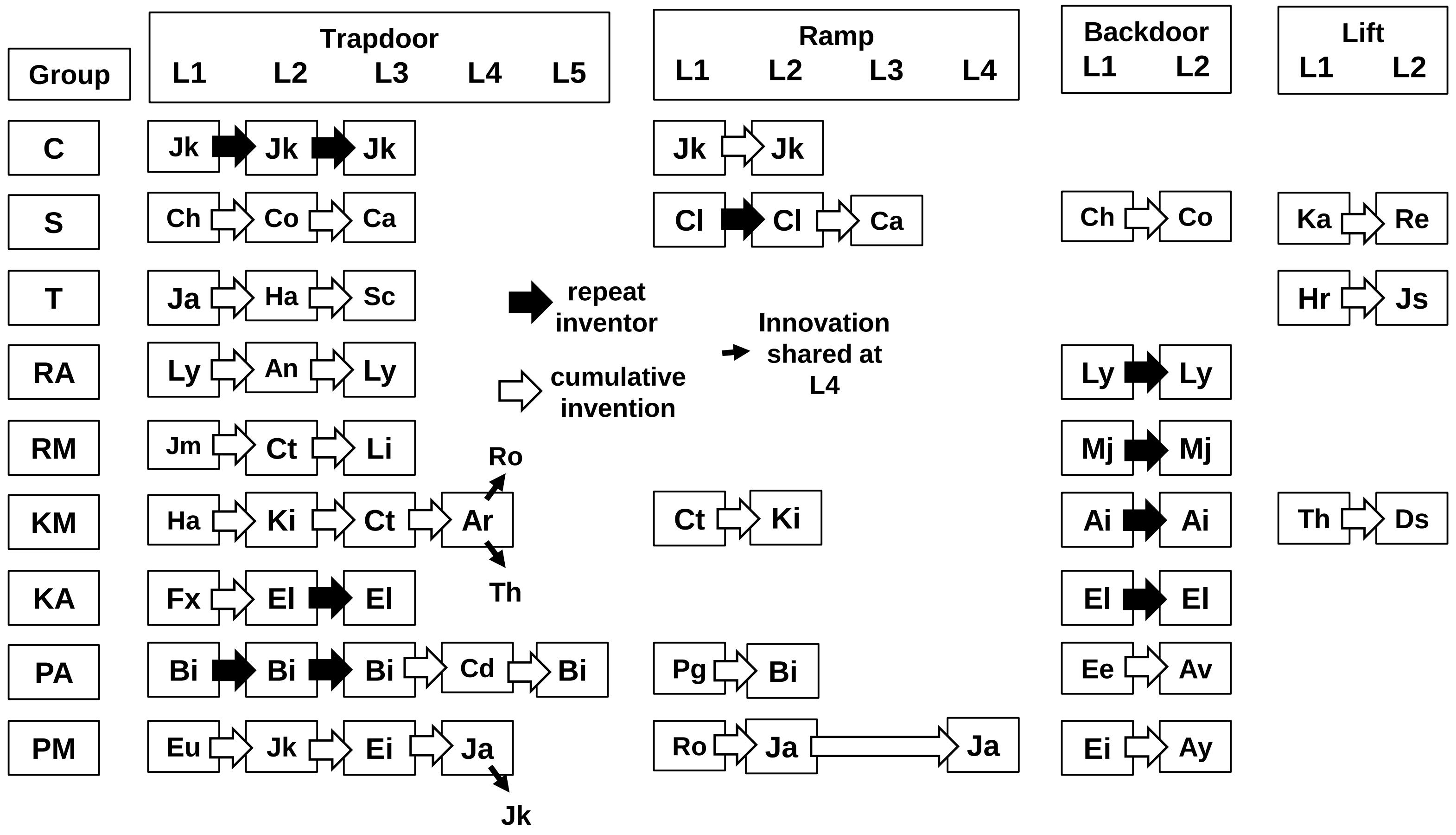
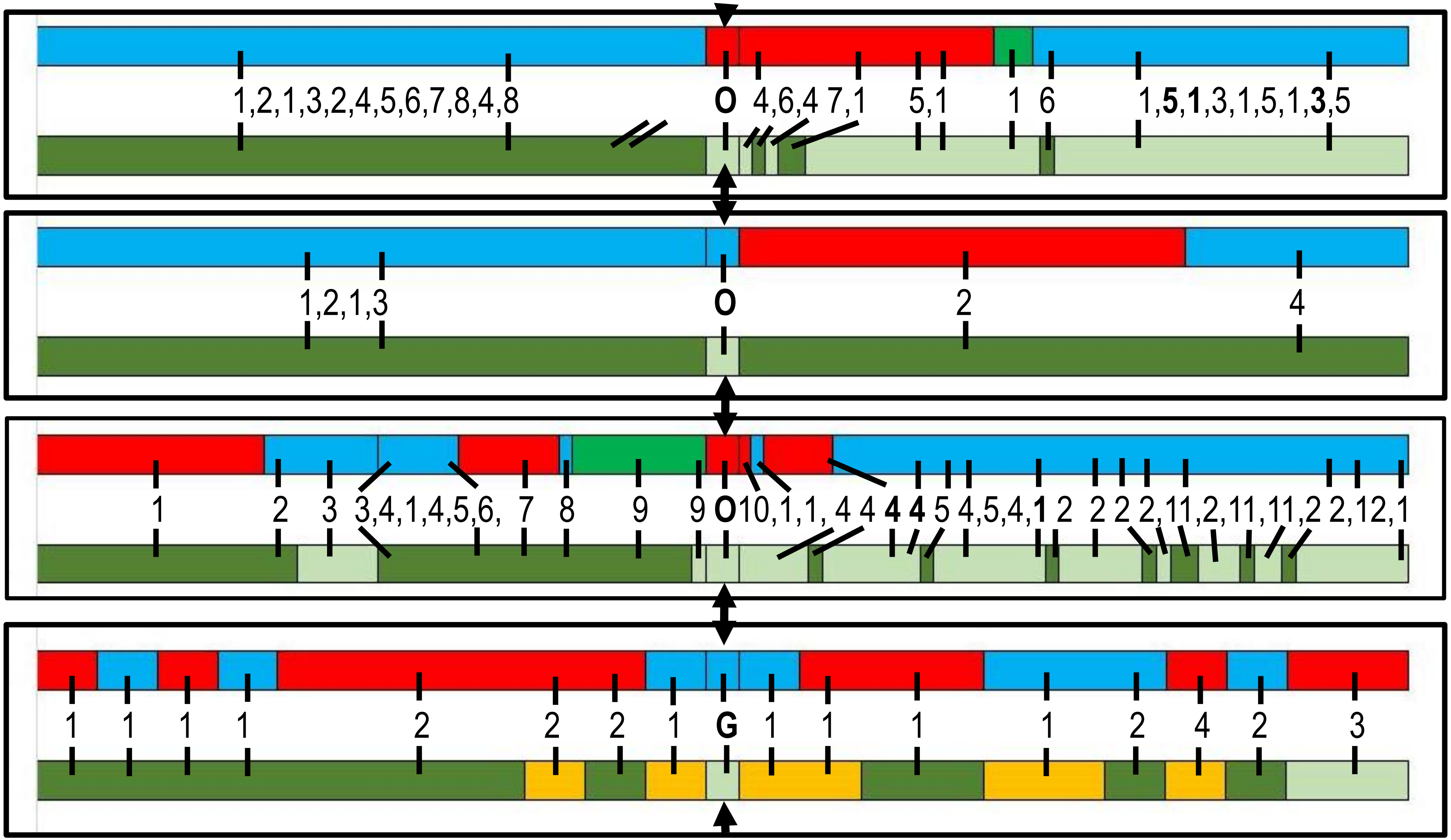

1
0.8
0.6
0.4
0.2
0
Ramp
Trapdoor
Pushdoor
Lift
*
1
0.8
0.6
0.4
0.2
0
Ramp
Trapdoor
Pushdoor
Lift
*
1
0.8
0.6
0.4
0.2
0
Trapdoor
*

Figure 7.



Group
Trapdoor
L1 L2 L3 L4 L5
Ramp
L1 L2 L3 L4
Backdoor
L1 L2
Lift
L1 L2
C
Jk Jk Jk
Jk Jk
S
Ch Co Ca
Cl Cl Ca
Ch Co
Ka Re
T
Ja Ha Sc
Hr Js
repeat inventor
cumulative invention
Innovation shared at L4
RA
Ly An Ly
Ly Ly
RM
Jm Ct Li
Mj Mj
Ro
KM
Ha Ki Ct Ar
Ct Ki
Ai Ai
Th Ds
Th
KA
Fx El El
El El
PA
Bi Bi Bi Cd Bi
Pg Bi
Ee Av
PM
Eu Jk Ei Ja
Ro Ja Ja
Ei Ay
Jk

Figure 8.



KM
1,2,1,3,2,4,5,6,7,8,4,8
O
4,6,4 7,1
5,1
1 6
1,5,1,3,1,5,1,3,5
KA
1,2,1,3
O
2
4
PM
1
2
3
3,4,1,4,5,6,
7
8
9
9
O
10,1,1, 4 4 4 4 5 4,5,4,1 2 2 2 2,11,2,11,11,2 2,12,1
PA
1 1 1 1 2 2 2 1 G 1 1 1 1 2 4 2 3